

**Economic and Social Council**

Distr.: General
12 April 2017

Original: English

For decision

United Nations Children's Fund

Executive Board

Annual session 2017

13-16 June 2016

Item 7 of the provisional agenda*

**Annual report for 2016 on the evaluation function
in UNICEF***Summary*

The present report documents evaluation activities undertaken by UNICEF in 2016. It provides a review of the status of the evaluation function within UNICEF and a report on progress made in the implementation of the revised Evaluation Policy (E/ICEF/2013/14). The report also covers the governance of the function within UNICEF and efforts made to promote evaluation coherence within the United Nations system and more widely. It presents updates on performance and results, a summary of the findings from recent assessments of the evaluation function and an annex detailing the status of evaluations proposed in the plan for global thematic evaluations for 2016. Elements of a decision for consideration by the Executive Board are also provided.

* E/ICEF/2017/4

17-06083 (E) 190417



Please recycle 



I. Introduction

1. Towards the end of 2016, the General Assembly selected Antonio Guterres as the new Secretary-General of the United Nations. Before his appointment, he set out his vision for the Organization in a comprehensive statement, which emphasized the importance of transparency, accountability and oversight. Speaking to delegates in April 2016, he further observed that, in the United Nations, "We need a culture of evaluation — independent and real-time evaluation with full transparency."¹

2. UNICEF is committed to transparency, accountability and effective oversight, and the evaluation function in UNICEF contributes towards this. In line with its revised Evaluation Policy (E/ICEF/2013/14), UNICEF uses evaluation to demonstrate results, promote accountability and provide evidence for better policies and programmes. In this way, evaluation contributes directly towards the achievement of the mission, mandate and priorities of the organization and, more widely, of the United Nations system.

3. The present report documents the evaluation activities undertaken by UNICEF in 2016. It reports on progress in the implementation of the revised Evaluation Policy and presents updates on performance and results. As in previous years, the report considers the governance of the function within UNICEF and the efforts made to promote evaluation coherence within the United Nations system and more widely. It also summarizes the findings of recent assessments of the evaluation function. Elements of a decision for consideration by the Executive Board are provided in section VIII.

II. Governance of the evaluation function

4. The Executive Board maintains oversight of the evaluation function within UNICEF and upholds the central role that it plays within the organization. At the annual session of the Executive Board in June 2016, the Evaluation Office presented the annual report for 2015 on the evaluation function, which was complemented by a management perspective presented by senior management. Responding to the annual report, Board members raised pertinent issues concerning the function and issued decision 2016/7, which addressed the following issues: (a) the geographical and thematic coverage of evaluations; (b) progress towards achieving the evaluation policy target of allocating a minimum of 1 per cent of overall programme expenditures to evaluation; (c) the timely delivery of management responses and their implementation; and (d) the development of national capacities for evaluation. At the Board's request, an action plan for achieving the 1-per-cent policy target was presented to the Board by management at the second regular session in 2016. Progress towards the target is described in section V.

5. The Executive Board also requested that an external peer review of the evaluation function be undertaken to meet a requirement in the revised Evaluation Policy. Arrangements for this are reported in section VI.

6. Evaluation reports, each accompanied by a management response, were presented at all three sessions of the Executive Board in 2016, with more extended presentations and discussions at informal briefings ahead of each session. The following documents were presented:

¹ United Nations Evaluation Group, "United Nations-designated Secretary-General calls for strong culture of evaluation" (www.unevaluation.org/mediacenter/newscenter/newsdetail/121).

(a) Plan for global thematic evaluations 2014-2017: review and update for 2016-2017 (E/ICEF/2016/3), which included the rescheduling of some evaluations listed in the original plan;

(b) "Protecting children from violence: A comprehensive evaluation of UNICEF's strategies and programme performance";

(c) "Equity, scalability and sustainability in UNICEF water, sanitation and hygiene (WASH) programming: Evidence from UNICEF evaluations 2007-2015, an evaluation synthesis report";

(d) "Increasing access and equity in early childhood education in the CEE/CIS region", an independent multi-country evaluation commissioned and managed by the Regional Office.

7. The Global Evaluation Committee, an internal body that provides oversight of evaluation issues, met in June 2016. Members approved changes in the membership and mode of operation of the Committee, considered proposals to strengthen evaluation coverage and the implementation of agreed actions in evaluation management responses and took note of arrangements for the external peer review.

8. The external Audit Advisory Committee considers evaluation to be an important element of the oversight system and monitors the performance of the evaluation function. In its 2015 annual report, presented to the Executive Board in 2016, the Committee took particular note of the need for management to continue to improve its follow-up to evaluation recommendations, and requested to be informed of progress in such follow-up.²

III. Promoting evaluation coherence within the United Nations

9. The quadrennial comprehensive policy review of operational activities for development of the United Nations system (QCPR),³ adopted by the General Assembly on 21 December 2016, provided an important orientation for all United Nations development activities, including evaluation. The QCPR emphasizes the need to better support the implementation of the 2030 Agenda for Sustainable Development through system-wide strategic planning, implementation, reporting and evaluation. In 2016, UNICEF continued to contribute towards the development of an independent system-wide evaluation mechanism.

10. In the 2016 QCPR, the General Assembly noted the need to strengthen support to national institutions in developing their capacities for planning, management and evaluation. UNICEF support to national evaluation capacity development in 2016 is described in section IV.

11. The General Assembly also urged organizations within the United Nations system to continue their efforts to harmonize their requirements regarding reporting, monitoring and evaluation in order to reduce transaction costs and boost efficiency. UNICEF recognizes the value of effective partnerships, and contributes to efforts to harmonize evaluation activities and build capacity, notably by developing agreed global standards, methods and agendas. Key initiatives include activities undertaken by the United Nations Evaluation Group (UNEG), joint evaluation activities and support to the independent system-wide evaluation mechanism.

² See www.unicef.org/about/execboard/files/UNICEF_Audit_Advisory_Committee_2015_AR-13Apr16-FINAL.pdf.

³ General Assembly resolution 71/243.

A. United Nations Evaluation Group

12. UNEG brings together evaluation offices across the United Nations system to undertake joint activities aimed at strengthening and harmonizing evaluation activities within the Organization and building stronger evaluation partnerships beyond the United Nations system. In 2016, UNICEF continued to play an active role in UNEG. The Director of the Evaluation Office continued to serve as UNEG Vice-Chair (Partnerships), one of four UNEG Vice-Chairs, and UNICEF staff participated in several working groups.

13. In 2016, UNEG concentrated its efforts to further professionalize the evaluation function. Key outputs included the publication of a revised and updated version of the UNEG Norms and Standards for Evaluation⁴ and the UNEG Evaluation Competency Framework.⁵ Evaluation Office staff contributed to task groups on evaluation use and professionalization, as well as to the UNEG Humanitarian Evaluation Interest Group, which, in 2016, published a meta-analysis of evaluation practice entitled "Reflecting Humanitarian Principles in Evaluation".⁶

14. UNICEF also supported the advocacy of UNEG on evaluation issues, building upon the landmark General Assembly resolution 69/237, adopted in 2015, which called upon the United Nations to support Member States to build capacity for the evaluation of development activities at the country level. In the context of country-led activities geared towards the achievement of the Sustainable Development Goals, support for enhancing the capacity of national evaluation systems takes on particular significance. In this regard, working with sister agencies, UNICEF hosted two well-attended events in 2016 focused on evaluation and the Goals, featuring high-level speakers from a wide range of countries as well as the United Nations. The events were arranged as side meetings of, respectively, the sixtieth session of the Commission on the Status of Women⁷ and the meeting of the high-level political forum on sustainable development, held under the auspices of the Economic and Social Council.⁸

B. Independent system-wide evaluation of United Nations operational activities for development

15. In the 2012 QCPR (resolution 67/226), the General Assembly provided for the establishment of an independent system-wide evaluation mechanism and reiterated in the 2016 QCPR its call for such a mechanism, specifically underscoring "the importance of strengthening a high-quality, independent and impartial system-wide evaluation mechanism of operational activities for development that enhances coherence and interdependence in the overall evaluation architecture of the United Nations development system ... using its findings and recommendations to improve the functioning of the system".

16. UNICEF has supported work on the mechanism from the outset, contributing funding and making significant technical contributions through participation in evaluation management groups established to pilot two independent, system-wide evaluations. In 2016, two evaluations were successfully completed:

⁴ www.uneval.org/document/detail/1914.

⁵ www.uneval.org/document/detail/1915.

⁶ www.uneval.org/document/detail/1862.

⁷ www.evalpartners.org/evalgender/evaluating-SDGs-with-an-equity-focused-and-gender-responsive-lens.

⁸ <https://sustainabledevelopment.un.org/index.php?page=view&type=20000&nt=480&menu=2993>.

(a) "Meta-evaluation and synthesis of United Nations Development Assistance Framework evaluations, with a particular focus on poverty eradication".⁹

The evaluation findings called for improved coverage, quality and resourcing of UNDAF evaluations;

(b) "Evaluation of the contribution of the United Nations Development system to strengthening national capacities for statistical analysis and data collection to support the achievement of the Millennium Development Goals and other internationally agreed development goals".¹⁰

In view of the 2030 Agenda for Sustainable Development, the evaluation report called for the strengthening of joint data collection and statistical capacity development across the United Nations system.

17. The independent system-wide evaluation mechanism was independently reviewed at the end of 2016 and the report of the review is forthcoming.

C. Inter-agency Humanitarian Evaluation Steering Group

18. The Inter-agency Humanitarian Evaluation Steering Group, which comprises both United Nations agencies and humanitarian non-governmental organizations, has continued to undertake joint evaluation activities. In 2016, the group disseminated the findings from the inter-agency evaluation of the United Nations response to the crisis in South Sudan¹¹ (completed in 2015). It also completed three additional reports:

(a) "Syria coordinated accountability and lesson learning (CALL) evaluation synthesis and gap analysis",¹² which covers the main issues highlighted in 24 publicly available evaluations and evaluative studies concerning the international response to the Syria crisis between 2012 and 2015;

(b) "Inter-agency humanitarian evaluation of the response to the crisis in the Central African Republic",¹³ which concluded that the inter-agency response made a large positive impact on the crisis, although coverage remained a major challenge;

(c) "Synthesis of key findings from inter-agency humanitarian evaluations of the international responses to crises in the Philippines (Typhoon Haiyan), South Sudan and the Central African Republic".¹⁴

This synthesis of the key findings drew out a wide range of lessons relating to needs assessment and prioritization; strategic and operational planning; measuring effectiveness; leadership and accountability; stakeholder engagement; and questions concerning the need for context-sensitive application of the Humanitarian Programme Cycle.

IV. Promoting evaluation coherence globally

A. Development networks

⁹ www.unjui.org/en/reports-notes/JIU%20Products/JIU_REP_2016_6_Final_English.pdf.

¹⁰ www.unjui.org/en/reports-notes/JIU%20Products/JIU_REP_2016_5_Final_English.pdf.

¹¹ www.unicef.org/evaldatabase/index_92844.html.

¹² www.interagencystandingcommittee.org/system/files/ocha_syria_web_interactive_2.pdf.

¹³ www.interagencystandingcommittee.org/system/files/print_final.pdf.

¹⁴ www.interagencystandingcommittee.org/system/files/web_interactive_0.pdf.

19. The main channel for UNICEF engagement in the promotion of evaluation coherence at the global level beyond the United Nations networks has been through EvalPartners, a global movement aimed at enhancing the capacities of civil society organizations to influence policymakers and to engage in national evaluation processes, through innovative partnerships and networking.¹⁵ EvalPartners is a partnership that links UNEG with the International Organization for Cooperation in Evaluation, a network of national and regional evaluation associations. UNICEF co-founded EvalPartners in 2012; the partnership has since become increasingly active and influential. A key output has been the development, through a consultative process, of the Global Evaluation Agenda 2016-2020, which is focused on evaluation capacity development and professionalization. The publication and dissemination of the Global Evaluation Agenda was supported by UNICEF. The organization also contributed substantially to the fulfilment of EvalPartners' mission and objectives in 2016. Throughout the year, the Director of the UNICEF Evaluation Office served as Co-Chair of EvalPartners, in his capacity as UNEG Vice-Chair (Partnerships).

20. EvalPartners has generated a set of global networks that operate under its umbrella. UNICEF has played a leading role in EvalSDGs, a network formed to support the integration of evaluation into national, regional and global systems for the planning and review of progress towards the Sustainable Development Goals. Partnership with the International Institute for Environment and Development provided a channel for publishing a series of briefing papers on evaluation of the Goals. Five papers were produced in 2016, of which three were the most-downloaded publications of the Institute in 2016.

21. UNICEF also provides technical support to the Steering Committee of the Global Parliamentarians' Forum for Evaluation and, in 2016, contributed to the creation of a Eurasian Regional Parliamentarians' Forum for Evaluation. The UNICEF country office in Kyrgyzstan, working closely with the national authorities, played an important supporting role.

B. Humanitarian networks

22. In the field of humanitarian evaluation, UNICEF has continued to work with the Active Learning Network for Accountability and Performance in Humanitarian Action. In 2016, the network published the Evaluation of Humanitarian Action Guide. On the basis of the guide, UNICEF has developed a hands-on capacity-building initiative for UNICEF country offices and government partners on conducting evaluations (see section V), and also provided training for civil society organizations at the Second International Francophone Forum on Evaluation, held in Marrakesh, Morocco, in 2016.

V. Performance of the evaluation function

23. UNICEF reports on performance of the evaluation function against seven core indicators:

- (a) The number of evaluations managed and submitted to the Evaluation and Research Database;
- (b) Topical distribution;
- (c) Types of evaluations conducted;

¹⁵ See www.evalpartners.org.

- (d) Evaluation quality;
- (e) Evaluation use;
- (f) Corporate spending on evaluation;
- (g) Production of corporate evaluations.

Each indicator is discussed below. The overall picture that emerges is one of continued improvement, confirming the general picture presented in previous annual reports. However, this generally positive story is disturbed by variations in performance over time and by region. There is a clear need for greater consistency.

24. Key findings include the following:

(a) The number of evaluations submitted to the Evaluation and Research Database has returned to acceptable levels, reaching 90 evaluations in 2015 and 102 evaluations in 2016;

(b) However, the geographical coverage of UNICEF evaluation remains below 80 per cent of all countries in which UNICEF is active, and shows significant variation between regions;

(c) Thematic coverage shows a balanced topical distribution between child survival, education and child protection, but very low coverage of the cross-cutting issues of social inclusion, humanitarian action and gender equality;

(d) The quality of evaluations submitted in 2015 dipped significantly compared with previous years, but recovered in 2016. Analysis has linked the drop in quality to increased diversity in the evaluation portfolio;

(e) There was a significant improvement in the submission of management responses, with 95 per cent of evaluations submitted in 2015 followed by a management response. Attention still needs to be given to improving the implementation of management responses;

(f) Budget use for evaluation was more than \$50 million in 2016, an increase over 2015, when the corresponding figure was just under \$45 million. However, these figures represent only 0.8 per cent of the total budget use for each year. Further progress is required to meet the policy target of 1 per cent.

25. The indicators in this section draw upon data generated by the UNICEF Global Evaluation Reports Oversight System (GEROS). Evaluations submitted to the Evaluation and Research Database are routinely assessed by an external company and a meta-evaluation of results is published each year. Evaluations submitted to the database in 2015 were assessed and rated for quality in 2016 and the results were published.¹⁶

Indicator 1: Number of evaluations managed and submitted to the Evaluation and Research Database

26. In its decision 2016/7,¹⁷ the Executive Board requested a geographic and thematic overview of the coverage of evaluations in the next annual report on the evaluation function in UNICEF. The overview is provided in this segment, on indicator 1, and in the segment on indicator 2, topical distribution.

¹⁶ "UNICEF Global Evaluation Reports Oversight System (GEROS) meta-analysis 2015: An independent review of UNICEF evaluation report quality and trends, 2009-2015" (https://www.unicef.org/evaldatabase/index_92693.html).

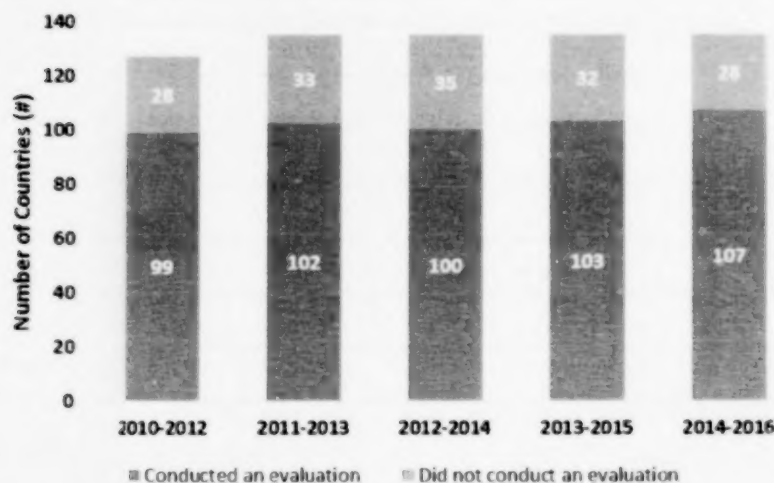
¹⁷ See E/ICEF/2016/15.

27. In 2015, 90 evaluation reports were submitted to the Evaluation and Research Database, an increase of 13 per cent over the total of 80 evaluations submitted in 2014. The number of evaluation reports submitted in 2016 indicates a continued increase, with 102 reports.

28. Despite fluctuations in the total number of reports submitted, geographical coverage has changed little over the period 2010-2015, as shown in figure 1. Here, geographical coverage is considered to be the proportion of country offices that submitted at least one evaluation report to the database in the reporting year or in the preceding two years. Geographical coverage remained at approximately 75 per cent over the period. In the period 2013-2015, 103 countries (76 per cent) submitted at least one evaluation report, while 32 country offices (24 per cent) did not. However, 9 of the 32 offices submitted evaluation reports in 2016.

29. In view of the last point, it might be expected that global coverage in 2016 would improve. An analysis of geographical coverage in the 2014-2016 period indicates that while 107 country offices submitted evaluations in the relevant period — an improvement on previous years — as many as 28 country offices did not. The global coverage rate for 2016 is therefore similar to previous years and remains below 80 per cent.

Figure 1
Evaluation coverage trends, 2010-2016



Source: UNICEF Global Evaluation Reports Oversight System (GEROS).

30. Regional analysis throws light on the issues. The number of evaluations submitted in 2015 varies by region, with the Central and Eastern Europe and the Commonwealth of Independent States (CEE/CIS), East Asia and the Pacific (EAPR), Eastern and Southern Africa (ESAR), South Asia (SA) and West and Central Africa (WCAR) each submitting 12 or more evaluations, while the other regions submitted fewer (see table 1).

Table 1
Regional evaluation submission rates, 2012-2016

Year	Region							Total
	CEE/CIS	EAPR	ESAR	LACR	MENA	SA	WCAR	
2012	16	8	29	10	10	7	34	114
2013	17	7	36	17	7	4	24	112
2014	15	6	33	8	4	14	12	72
2015	14	14	17	7	8	12	15	87
2016	9	12	26	13	6	13	13	92

Abbreviations: Central and Eastern Europe and Commonwealth of Independent States (CEE/CIS), East Asia and the Pacific (EAPR), Eastern and Southern Africa (ESAR), Latin America and Caribbean (LACR), Middle East and North Africa (MENA), South Asia (SA) and West and Central Africa (WCAR).

Source: GERO.

31. It is more meaningful to consider the coverage rates and trends by region, however. In the period 2013-2015, all countries in SA submitted at least one evaluation, resulting in a regional geographical coverage rate of 100 per cent. ESAR was close behind, with coverage of 95 per cent, and CEE/CIS performance was also above average, with 82 per cent coverage, while other regions were trailing. Figures for the period 2014-2016 show greater consistency across regions (see table 2). The sharp improvement in coverage in EAPR is especially noteworthy.

Table 2
Geographical evaluation coverage by region, 2013-2016

Period	Region (percentage)						
	CEE/CIS	EAPR	ESAR	LACR	MENA	SA	WCAR
2013-2015	82	73	95	56	65	100	76
2014-2016	86	93	82	64	65	100	80

Source: GERO.

32. Evaluation coverage is dependent upon the conditions prevailing in the different countries and regions and the resources available to regional and country offices. The countries that struggle to undertake evaluations tend to be those with small programmes and limited resources. Creative evaluation planning has successfully addressed this issue in CEE/CIS, where the Regional Office has for several years played a key role in coordinating and managing multi-country evaluations, drawing on pooled resources. This approach could be applied elsewhere.

33. The political context poses a major constraint to conducting evaluations in some countries, with considerable challenges in countries experiencing conflict. Yet in the Middle East and North Africa region (MENA), for example, several UNICEF offices have found ways of undertaking useful evaluations despite protracted armed conflict and the immense challenges affecting the region.

34. Finally, an analysis of regional productivity over time yields some surprises. In two regions, as shown in table 1, the number of evaluations completed has doubled in comparison with earlier years. Production in SA jumped from 4 evaluations in 2013 to 14 in 2014, and in EAPR, the number of evaluations

increased from 6 in 2014 to 14 in 2015. In other regions, the numbers dropped by half. Both ESAR and WCAR submitted a high number of evaluations in 2012 and 2013, but output dropped by around 50 per cent in the two succeeding years. The cases of high productivity demonstrate what is possible, while pervasive fluctuations indicate the importance of sustained effort and close oversight.

Indicator 2: Topical distribution

35. This segment is concerned with thematic coverage, an issue on which the Executive Board has requested information. Thematic coverage is addressed through annual reporting against the key performance indicator of topical distribution, which reviews the evaluation portfolio in terms of the various outcome areas and cross-cutting issues, or themes, set out in the UNICEF Strategic Plan, 2014-2017.

36. For the present analysis, 87 reports submitted in 2015 were considered.¹⁸ The results show a balanced topical distribution between child survival, education and child protection, with child survival (comprising health, HIV/AIDS, nutrition and WASH) the subject of 25 reports (29 per cent); education the subject of 24 reports (28 per cent); and child protection the subject of 22 reports (25 per cent).¹⁹

37. Only a handful of the 87 evaluations reviewed were focused specifically on the cross-cutting themes of social inclusion, humanitarian action and gender, despite the importance of those areas. Social inclusion is a key outcome area under the Strategic Plan, 2014-2017, but only two evaluation reports on that topic were submitted. Although coverage of these issues appears to be low, many evaluations that are focused mainly on sectoral issues do include some coverage of social inclusion, humanitarian action and gender equality.

38. Humanitarian action is included in the Strategic Plan as an important cross-cutting issue, yet only two evaluation reports that were focused specifically on humanitarian issues were submitted. In recent years, humanitarian action has been a major and growing area of work for UNICEF. While UNICEF has undertaken various unpublished reviews and lesson-learning exercises, there is a need across UNICEF for the systematic evaluation of humanitarian action and the necessary follow-up.

39. Recognizing that limited evaluation "know-how" can be a constraint in humanitarian emergencies, the Evaluation Office has launched a programme of hands-on training and follow-up support to assist country offices in undertaking evaluations. The programme is aimed not only at developing the capacity of country offices to evaluate humanitarian action but also to produce a set of evaluations. It is expected that the programme will boost coverage in the coming years.

40. In 2016, the programme supported the evaluation of humanitarian activities in 17 countries, including some in very challenging situations. Three of these evaluations were completed and submitted in 2016, with a dozen more expected to be reported in 2017. In addition, the Evaluation Office is preparing a synthesis of available evaluations of humanitarian response, due for submission to the Executive Board in September 2017.

41. Although gender equality is an important cross-cutting issue under the Strategic Plan, it was the specific focus of only one evaluation report reviewed. However, it can be noted that, between 2012 and 2015, there was an increase in the share of evaluations of sectoral topics that also addressed gender equality, equity

¹⁸ Three reports could not be classified under a specific topic and were excluded from the analysis.

¹⁹ UNICEF Geros meta-analysis 2015.

and human rights as key themes. The “United Nations Children’s Fund (UNICEF): Review of development effectiveness, 2012-2015”, commissioned in 2016, also provides a strong indication that UNICEF has responded appropriately by refocusing evaluation attention to these themes.²⁰ The review shows that, compared with a similar exercise undertaken in 2012, there was an increase in the share of evaluations that addressed gender equality (58 of 66), equity (54 of 66) and human rights-based approaches to programming (39 of 66) as key elements. Inclusion of these themes in the majority of the evaluations has arguably reduced the need for evaluations specifically focused on them.

42. Analysis of the topical distribution over time shows some significant shifts. In 2015, there was an increase in the percentage of sector-specific evaluations, with 84 per cent of the evaluations focused on a specific sector. This was an increase from 66 per cent and 41 per cent reported in 2014 and 2012, respectively. Conversely, as noted above, there was a decline in the share of evaluation reports that specifically addressed cross-cutting issues: from 9 per cent in 2014 to only 5 per cent in 2015. One remarkable feature was the increase in the percentage of evaluation reports focused on child protection, which rose from 6 per cent in 2014 to 25 per cent in 2015.

Table 3
Topical distribution of evaluation reports, 2010-2015

Topic	Percentage			
	2010	2012	2014	2015
Sector specific, by key results area of the Strategic Plan, 2014-2017	66	41	66	84
Child survival and development		26	23	22
Health			13	15
Nutrition			0	1
Water, sanitation and hygiene			10	6
Education	15	13	32	28
Child protection	15	8	6	25
HIV/AIDS	10	5	4	7
Social inclusion	1	1	1	2
Programmes covering several sectors	34	43	16	13
Cross-cutting themes	–	10	9	5
Organizational performance/other	–	6	9	3

Source: GEROs.

Indicator 3: Types of evaluation conducted

43. Analysis of evaluation reports submitted in 2015 in terms of the results level assessed presents a clear focus on higher-level results. The great majority of evaluations were pitched at the outcome level (61 per cent) and the impact level (29 per cent) rather than the output level (10 per cent). This focus on higher-level results indicates an appropriate use of evaluation in understanding UNICEF contributions and achievements.

44. Further analysis (see table 4) reveals that the majority of outcome-level evaluations were rated “highly satisfactory” or “outstanding” and none were rated

²⁰ www.unicef.org/evaldatabase/index_94430.html.

"unsatisfactory". In contrast, well over half of the impact-level evaluations reviewed displayed some shortcomings. This suggests that UNICEF capacity is highest in the areas in which it has the most practice and experience, such as designing and implementing programmes and evaluations at the outcome level. This has significant implications for the overall quality of the evaluation portfolio, as examined in the segment on the quality of UNICEF evaluation reports.

Table 4
Evaluation focus and quality, 2015

Focus	Quality				Total
	Outstanding	Highly satisfactory	Mostly satisfactory	Unsatisfactory	
Impact		10	15	1	26
Outcome	3	31	21		55
Output	1	3	3	2	9
Total	4	44	39	3	90

Source: Geros.

45. Analysis of the evaluation reports submitted in 2015 indicates that most of the evaluations were managed directly by UNICEF offices. UNICEF-managed evaluations accounted for 72 per cent of the evaluation portfolio, while joint United Nations-managed, joint country-managed and country-led evaluations accounted for 18 per cent, 3 per cent and 1 per cent, respectively. The management of 6 per cent of the evaluation reports could not be ascertained.

46. Analysis of the evaluation reports submitted in 2015 indicates that the purpose of the evaluations was fairly evenly divided among formative evaluations (forward-looking and focused on learning), summative evaluations (backward-looking, usually with a strong focus on accountability) and evaluations that were both summative and formative (see table 5). This suggests, overall, an appropriate deployment of evaluation to support not only "real-time" learning, but also accountability and future improvement.

Table 5
Evaluation purpose, 2015

Purpose	Number	Percentage
Formative	23	26
Summative	30	33
Summative and formative	37	41

Source: Geros.

Indicator 4: Quality of UNICEF evaluation reports

47. The following findings are based on the quantitative and qualitative analysis of data from the Geros Evaluation Quality Assurance and United Nations System-wide Action Plan on Gender Equality and the Empowerment of Women (SWAP) reviews of 90 UNICEF evaluation reports submitted to the Evaluation and Research Database in 2015.

48. In 2015, the majority of evaluation reports were almost equally divided between the ratings of "highly satisfactory" (49 per cent) and "mostly satisfactory" (43 per cent) (see table 6). Four reports (4 per cent) were rated "outstanding" and three "unsatisfactory". In other words, 53 per cent of the reports submitted were deemed "good", i.e., "highly satisfactory" or better, while only 3 per cent were deemed "unsatisfactory". Nevertheless, 43 per cent of the evaluations assessed showed some shortcomings and further analysis, below, explores why this might be so.

49. The quality of evaluation reports varies considerably by sector. Most of the evaluation reports on humanitarian action and WASH were found to be "highly satisfactory" or better (although with small sample sizes). The majority of evaluation reports in areas concerned with child protection, HIV/AIDS and social inclusion as well as multisector evaluations were also found to be "highly satisfactory". However, less than half of the 24 evaluation reports in the education sector were rated "satisfactory", and only 4 of the 13 evaluation reports in the health sector were rated "satisfactory".

50. Overall, the ratings indicate a sharp decline in the quality of evaluation reports submitted in 2015 compared with those submitted in 2014. While 74 per cent of UNICEF evaluations submitted in 2014 were rated as "highly satisfactory" or "outstanding", only 53 per cent of the 2015 submissions achieved such ratings. At the same time, evaluation reports rated as "mostly satisfactory" increased from 23 per cent in 2014 to 43 per cent in 2015, while the percentage of evaluations rated as "unsatisfactory" remained low, at 3 per cent.

51. Results for evaluations submitted in 2016 indicate a much more positive picture, however. Among the 101 evaluations rated, 74 per cent are judged to be "highly satisfactory" or "outstanding" and only one evaluation is rated as "unsatisfactory". This indicates that the drop in quality observed in evaluations submitted in 2015 was an anomaly, and that the improving trend reported in previous years has resumed.

Table 6
Quality of completed evaluations, 2009-2015, by percentage

Quality rating	Year				
	2009	2012	2014	2015	2016
Outstanding	4	3	6	3	6
Highly satisfactory	32	59	68	49	68
Mostly satisfactory	34	30	23	43	25
Unsatisfactory	30	8	3	3	1

Source: GEROS.

52. Nevertheless, the dip in the quality of the 2015 evaluation reports was significant and calls for explanation. Further, to reverse this decline, it is important to understand not only the likely causes, but also to verify whether the decline actually occurred or was an effect of inconsistent measurement.

53. The decline in quality in 2015 coincided with a change in the external firm that provides the quality assurance on UNICEF evaluation reports. This raises the question of whether the standard of assessment has become more rigorous, resulting in stricter ratings. The Evaluation Office took steps to manage this risk by providing intensive training to the new company so that the same standards would be applied and by the close monitoring of the quality review process. The Evaluation Office is confident that consistent standards were applied.

54. The meta-analysis of the ratings reveals a more significant issue. The percentage of reports rated as meeting UNICEF standards appears to be proportional to the percentage of reports of outcome-level evaluations. Among evaluation reports submitted in 2015, 62 per cent of reports on outcome evaluations were rated as "highly satisfactory" or better, while only 38 per cent of reports on impact evaluations were so rated. Further, the proportion of reports concerned with impact evaluations was higher than in previous years. Taken together, these factors seem to have contributed to the overall reduction in quality noted for 2015. The implication is that a more diverse portfolio of evaluation types makes it more difficult to consistently deliver reports that meet the required quality standards. This further implies that additional support is required to ensure that all types of evaluations (and the associated evaluation reports) reach the required level of quality. Further monitoring is required to confirm this hypothesis.

55. Additional support for the proposition that it can be difficult to sustain quality in the portfolio when addressing complex topics or applying advanced methods derives from an examination of the evaluations receiving lower ratings. Among the evaluations submitted in 2015 that were rated "mostly satisfactory" appears a cluster of six evaluations of health systems strengthening, a new and challenging area of work for UNICEF. This points to the conclusion that, while it is important to maintain quality, it is also important to allow space for innovation and learning. Rating and reporting are important mechanisms for supporting learning. Accordingly, for each quality assessment of an evaluation, executive feedback is provided to the office submitting the evaluation.

United Nations System-wide Action Plan on Gender Equality and the Empowerment of Women

56. UNICEF monitors and reports on its evaluation performance against key indicators set out in the SWAP. For evaluations submitted in 2015, the aggregated average score for UNICEF was 6.36, which is classified as "approaching requirements". This represents a year-on-year improvement, with UNICEF reporting a rating of 6, "approaching requirements", in the 2014 cycle. This rating is consistent with that of similar United Nations entities, including the United Nations Development Programme, UN-Women and the United Nations Educational, Scientific and Cultural Organization.

57. Reports were found to be "satisfactory" with regard to integrating gender in the scope, indicators, criteria and questions of evaluations, but the use of gender analysis to inform evaluation findings, conclusions and recommendations was rated lower, and needs improvement.

Indicator 5: Use of evaluation, including management responses

58. The management response is a key feature of the evaluation cycle. In its decision 2016/7, the Executive Board urged UNICEF to prioritize the timely delivery of management responses and to improve the rate of implementation of agreed evaluation recommendations. In response, UNICEF has taken action, in line with the evaluation policy, to ensure that all evaluation reports that are uploaded into the Evaluation and Research Database are accompanied with a management response. Of evaluation reports submitted in 2015, 86 reports (95 per cent) were followed by a management response, an increase of 17 per cent over the 79 per cent reported for evaluations submitted in 2014.

59. Management response actions are supposed to be completed within one year. For evaluations submitted in 2015, 76 per cent of the agreed actions set out in management responses were reported to have been completed (33 per cent) or under

implementation (43 per cent) by the end of 2016. This repeats the rate achieved for evaluations submitted in 2014.

60. Analysis by region reveals considerable variation in the completion and implementation of agreed actions. For evaluations submitted in 2015, MENA has reported good progress, with 85 per cent completion of agreed actions, 9 per cent under implementation and only 6 per cent not started. ESAR has also made good progress, reporting 62 per cent of agreed actions completed, 33 per cent under implementation and only 6 per cent not started. EAPR has also reported above-average performance (56 per cent completion, 31 per cent under way). However, other regions are lagging to varying degrees. Once again, good performance in some regions demonstrates what is possible, while flagging the need for greater consistency across regions.

Indicator 6: Corporate spending on evaluation

61. The UNICEF evaluation policy sets a target for spending on evaluation of 1 per cent of programme expenditure. The Executive Board has been monitoring this closely and has urged UNICEF to accelerate progress towards achieving that target.

62. At the end of 2016, evaluation budget use was reported to have reached \$50,121,716, representing 0.8 per cent of total programme budget use.²¹ The corresponding figure at the end of 2015 was \$44,959,817, which again represented 0.8 per cent of total programme budget use. While these figures show a significant increase in budget use for evaluation, there was evidently no progress towards the policy target of 1 per cent over the year.

63. The figures are drawn from VISION, the UNICEF enterprise resource management platform. In 2016, the method used to calculate evaluation budget use was reviewed, revealing that important areas of expenditure were being omitted. Consequently, the Evaluation Office revised the methodology to include all expenditure on evaluation that can be readily tracked in UNICEF systems.²² The same methodology was used to calculate the figures for 2015 and 2016 presented above.

Indicator 7: Corporate-level evaluations

64. Activities considered under this indicator are corporate-level evaluations undertaken by the Evaluation Office. These include evaluations listed under the plan for global thematic evaluations 2014-2017 (E/ICEF/2014/3) and the review and update for 2016-2017 (E/ICEF/2016/3).

65. In 2016, the Evaluation Office completed and published the following evaluations and related studies:

- (a) Global thematic evaluations:²³
 - (i) "Evaluation of UNICEF's PMTCT/Paediatric HIV care and treatment programme";²⁴

²¹ Budget use, as recorded in VISION, the UNICEF enterprise resource management platform, includes actual expenditure and spending commitments.

²² Previously, evaluation budget use estimation was computed against regular resources and other resources — regular. For a more complete estimation of evaluation budget use, the calculation now includes evaluation costs recorded in several other budget lines.

²³ Global thematic evaluations are wide-ranging exercises intended to take a comprehensive view of the work of UNICEF in a particular subject area or theme and to provide, in each case, findings, conclusions and recommendations of strategic relevance to the organization.

²⁴ www.unicef.org/evaluation/index_95015.html.

- (ii) "Communication for development: An evaluation of UNICEF's capacity and action",²⁵
 - (b) Evaluations of UNICEF humanitarian action:
 - (i) "Evaluation of UNICEF's response to the Ebola outbreak in West Africa 2014-2015";²⁶
 - (ii) "The UNICEF response to the crisis in the Central African Republic";²⁷
 - (c) Evaluation synthesis studies, based on meta-analysis of existing evaluations:
 - (i) "Equity, sustainability and scalability in UNICEF WASH programming: Evidence from UNICEF evaluations 2007-2015";²⁸
 - (ii) "United Nations Children's Fund (UNICEF): Review of development effectiveness, 2012-2015";²⁹
 - (d) Country reports on developmental evaluation³⁰ activities of the UNICEF Peacebuilding, Education and Advocacy Programme in Ethiopia³¹ and Myanmar.³²
- Senior management has prepared management responses to the recommendations presented by these evaluations.
66. The Evaluation Office also participated in the management of several joint evaluations, as noted elsewhere in the present report. In collaboration with the Evaluation Office of the United Nations Population Fund, work continued on a joint health sector evaluation: End Line Evaluation of the H4+ partnership joint support to improve women's and children's health 2011-2016.
67. Progress in delivering the evaluations listed in the plan for global thematic evaluations is outlined in the annex.

VI. Assessments of the evaluation function

68. Section VI provides information on several assessments of the evaluation function in UNICEF, which have been completed. In addition, an external peer review is under way.

A. United Nations Evaluation Group/Developmental Assistance Committee peer review

69. The Executive Board has requested that an external peer review of the evaluation function be conducted, in line with a provision in the revised evaluation policy. The Organization for Economic Development-Developmental Assistance Committee (OECD-DAC), the Network on Development Evaluation (EvalNet) and UNEG maintain a joint task force to support professional peer reviews of the evaluation function in United Nations organizations. Each peer review is intended to

²⁵ www.unicef.org/evaldatabase/index_C4D.html.

²⁶ www.unicef.org/evaldatabase/index_95016.html.

²⁷ www.unicef.org/evaldatabase/index_92843.html.

²⁸ www.unicef.org/evaldatabase/index_94431.html.

²⁹ www.unicef.org/evaldatabase/index_94430.html.

³⁰ Developmental evaluation is an approach to understanding the activities of an intervention operating in a dynamic, novel or complex environment.

³¹ www.unicef.org/evaldatabase/files/UNICEF-PBEA_LR_DE-Ethiopia.pdf.

³² www.unicef.org/myanmar/PBEA_Developmental_Evaluation_Myanmar.pdf.

identify good practices and opportunities to further strengthen the evaluation function in the agency under review, with a view to contributing to improved performance and results. UNICEF was one of the first agencies to be subject to peer review, in 2006, which led to the preparation of the first UNICEF evaluation policy. To date, 16 UNEG/DAC peer reviews have been conducted.

70. By the end of 2016, arrangements for a UNEG/DAC peer review were in place, to be taken forward by a panel chaired by the Evaluation Director of the United Nations Environment Programme. Terms of reference were agreed for an assessment of the independence, credibility and utility of the evaluation function in UNICEF. These were focused on the independence and credibility of the function; the quality, use and follow-up of evaluations across the organization to promote accountability, learning and improvement; and the leadership and vision shown by the UNICEF evaluation function, including its capacity to adjust to the changing environment.

71. Arrangements were made for the panel to visit several UNICEF locations in 2017, with a view to delivering a report to UNICEF management and the Executive Board in the course of the year. It is noteworthy that this peer review will be the first to use the revised UNEG evaluation norms and standards, approved in 2016.

B. Self-assessment of the UNICEF evaluation function

72. In 2016, the Evaluation Office initiated a self-assessment of the evaluation function in UNICEF, with the purpose of generating evidence to inform the scheduled external peer review; assessing how well UNICEF was responding to ongoing efforts among UNEG members to increased professionalism; and informing a possible future revision of the UNICEF evaluation policy.

73. The self-assessment, coordinated by the Evaluation Office, was underpinned by a methodology that included the collation and review of relevant documentation, several online surveys and various participatory consultations. Recalling that evaluation is a shared function, particular attention was given not only to the roles and perceptions of monitoring and evaluation staff across UNICEF, but also to the views of users of evaluation evidence, including senior management.

74. Shadowing the approach of the external peer review, the self-assessment focused on the UNEG norms and standards, especially those relating to the independence, credibility and utility of the function. The results were widely shared with staff and management and with the incoming external peer review panel.

75. Key findings include the following:

(a) The evaluation function in UNICEF is perceived as credible by staff and management across the organization;

(b) The revised evaluation policy is widely known and accepted as the framework guiding evaluation practice in UNICEF. However, some aspects could be adjusted and updated;

(c) The highly decentralized evaluation model is perceived as an asset for the organization. However, the three-tier structure poses challenges for the coherent management of the function;

(d) While the evaluation function is considered to enjoy a reasonably high level of independence, there are some issues related to independence that would benefit from further reflection. It is recognized that it is difficult to maintain separate responsibilities for evaluation in small country offices; that dependence on programme budgets may, without appropriate safeguards, put the independence of

evaluation efforts at risk; and that, at the corporate level, the appropriate separation of roles in sector teams is also needed;

(e) Evaluation is generally considered useful, but a more user-friendly and less technocratic and compliance-driven approach is needed. Timeliness, knowledge management and quality of recommendations are all areas for attention;

(f) There is widespread concern among monitoring and evaluation staff that they have insufficient time to devote to evaluation: in small offices, they report spending less than 15 per cent of their time on evaluation matters. In many cases, the same staff members are dealing not only with evaluation but also with performance monitoring and reporting and the monitoring of children's rights, as well as with research and, in some cases, social policy. While this is already unrealistic, growing demands relating to results-based management requirements add to the burden;

(g) Staff capacities vary and attention is needed to strengthen skills, knowledge and systems at all levels, including among users of evaluation evidence.

76. Overall, while the institutional environment for evaluation at UNICEF was found to be supportive and demand for evaluation was seen to be high, capacities and resources to respond to demand are constrained.

C. United Nations Children's Fund (UNICEF): Review of development effectiveness, 2012-2015³³

77. In 2016, the Evaluation Office commissioned a review to consider how effectively UNICEF had been achieving results for children, as assessed through careful and methodical scrutiny of a sample of 66 evaluations reporting in the period 2012-2015. The review included systematic attention to the contribution of monitoring and evaluation to the organization's performance and results.

78. The findings regarding the use of evaluation evidence to improve effectiveness were "strongly positive", with 80 per cent of the sample reporting findings assessed as "satisfactory" (21 per cent) or "highly satisfactory" (59 per cent) (see p. 47).

79. The review lists factors that contribute to the use of evaluation to improve effectiveness and presents the observation that evaluation reports "often noted...that lessons from the evaluation of a previous project or programme were used to strengthen the design of the current program" (p. 48). Several evaluations identified the following positive factors promoting the use of evaluation results to improve effectiveness:

- (a) Strong baseline collection;
- (b) Use of external evaluators and external evaluations to support learning;
- (c) Systematic approach to evaluation with a commitment to regular, ongoing evaluations and to covering most programmes;
- (d) Commitment to using the results of all types of evaluations (formative, midterm, end of programme or project) to improve programme design.

80. Negative factors were also identified from the evaluations reviewed. These factors "were focused on those projects and programmes which did not seem to be subject to regular evaluations" (p. 49), specifically:

³³ www.unicef.org/evaldatabase/index_94430.html.

(a) A primary focus on project monitoring with inadequate attention by UNICEF and the implementing partner to evaluation;

(b) The absence of any commitment to, or mention of, a regular programme of evaluation so that there is no culture of evaluation embedded in the programme;

(c) The renewal of programmes without a prior evaluation;

(d) A focus on reporting outputs and neglecting learning and accountability.

81. The review also assessed the effectiveness of systems and processes for evaluation in UNICEF. Here, 70 per cent of the evaluations reviewed reported positive findings on this issue. The review further notes that the evaluations reviewed in 2016 were more likely to report positive results on the use of evaluations to improve performance and results than those reviewed for a similar exercise in 2012.³⁴ The review observes that "evaluations reviewed in 2016 were of generally higher quality and suggests that UNICEF has been effective in improving the quality and utility of evaluations conducted in the past three years" (p. 47). Specific findings on gender are noted in section III of the present report.

82. These observations confirm the general picture of increasing effectiveness and utility and of an evaluation function in UNICEF broadly fit for purpose.

D. Multilateral Organization Performance Assessment Network UNICEF institutional assessment report 2015-2016

83. The Multilateral Organization Performance Assessment Network review provides a granular assessment of UNICEF, looking in fine detail at the agency's organizational and development effectiveness, including the evaluation function. Assessed against a set of microindicators, the review rates the evaluation function as "highly satisfactory" and several other factors as "satisfactory" or "highly satisfactory" (see table 7). However, UNICEF performance in the application of evaluative evidence was found to be weak.

Table 7
Evidence-based planning and programming applied

<i>No</i>	<i>Microindicator</i>	<i>Rating</i>
8.1	A corporate independent evaluation function exists	Highly satisfactory
8.2	Consistent, independent evaluation of results (coverage)	Satisfactory
8.3	Systems applied to ensure the quality of evaluations	Highly satisfactory
8.4	Mandatory demonstration of the evidence base to design new interventions	Unsatisfactory
8.5	Poorly performing interventions proactively identified, tracked and addressed	Highly satisfactory
8.6	Clear accountability system ensures responses and follow-up to and use of evaluation recommendations	Satisfactory
8.7	Uptake of lessons learned and best practices from evaluations	Satisfactory

Source: Multilateral Organization Performance Assessment Network report, annex I.

84. This weakness is reflected in the conclusion that "performance data is not yet always systematically used in planning and decision-making or to inform programmatic adjustments." Although it was recognized that "the quality of

³⁴ Organization for Economic Development-Developmental Assistance Committee. "Review of UNICEF's development effectiveness 2009-2011: Final report" (2013).

evaluations is improving overall, supported by robust quality assurance systems," it was also noted that there were "challenges in sustaining coverage levels" and that "the extraction and use of lessons from evaluations [was] not yet systematic or comprehensive" (p. 27).

85. Despite the positive ratings noted above and continuing improvements in the evaluation function, these judgments are borne out by the information presented elsewhere in the present annual report as well as in previous ones, and, usefully, point towards areas for improvement.

VII. Conclusion

86. The information presented in the present report confirms that, overall, UNICEF has an effective evaluation function that continues to develop and improve. The organization's highly decentralized structure presents practical management challenges: greater consistency in performance over time and across regions is needed. Further, as the development agenda becomes more complex and the role of UNICEF changes, there is a need to revisit methodological approaches and provide staff and development partners with appropriate tools and methods to conduct useful and insightful evaluations. A largely decentralized structure supports rapid, localized and informal learning, but more formal mechanisms for the uptake and application of evidence and lessons are also needed as UNICEF itself grows in size and complexity and engages with ever-wider stakeholder networks.

87. However, the recent reviews and assessments have reaffirmed that much of the strength of the evaluation function in UNICEF is the wide distribution of evaluation accountabilities and responsibilities across the organization and the engagement and enthusiasm of staff and management in efforts to deepen understanding of what works for children in a rapidly changing world.

VIII. Draft decision

The Executive Board,

Takes note of the annual report for 2016 on the evaluation function in UNICEF (E/ICEF/2017/11).

Annex

Status of evaluations proposed in the plan for global thematic evaluations for 2016

1. In 2014, the Executive Board endorsed the plan for global thematic evaluation 2014-2017 (E/ICEF/2014/3). It was revised and updated, with the revised plan endorsed by the Board in 2016 (E/ICEF/2016/3).

2. The table below sets out the status of evaluations planned for 2016. A list of all evaluations completed in 2016 appears in section V (indicator 7: corporate-level evaluations) of the present report.

Table 1
Topics proposed for global thematic evaluations, 2016

<i>Type</i>	<i>Topic</i>	<i>Status at end of 2016 and schedule for completion</i>
Major evaluation	Strengthening national health systems and capacity	Thorough scoping exercise completed and major evaluation launched Country case studies: Q4 2017 Comprehensive synthesis: Q3 2018
Major evaluation	Girls' education and gender equality	In progress Final report: Q2 2017
Major evaluation	Rural/small town water supply programming	In progress Final report: Q4 2017
Major evaluation	Preventing stunting: improving equitable use of nutritional support and care	In progress Final report: Q2 2017
Evaluations	Preventing mother-to-child transmission of HIV	Completed
Evaluation	Health: H4+ joint programme on maternal, newborn and child health (joint)	In progress Final report: Q2 2017
Evaluation synthesis	Synthesis of evaluation findings in the HIV and AIDS sector	In progress Final report: Q2 2017
Evaluation synthesis	United Nations Children's Fund (UNICEF): Review of development effectiveness, 2012-2015	Completed